ICIC Express Letters
Volume **11**, Number **7**, July **2017**

ICIC International ©2017 ISSN 1881-803X
pp. 1221–1228

HEARTBEAT MULTI-CLASSIFICATION ALGORITHM BASED ON HIERARCHICAL SUPPORT VECTOR MACHINE

Xin Liu[1], Yujuan Si[1,2], Di Wang[1] and Yanchun Liang[2,3]

[1]College of Communication Engineering
[3]College of Computer Science and Technology
Jilin University
No. 2699, Qianjin Street, Changchun 130012, P. R. China
liuxin52419558@126.com; { siyj; ycliang }@jlu.edu.cn; 15143086560m0@sina.cn

[2]Zhuhai College of Jilin University
Jinwancaotang, Zhuhai 519041, P. R. China

Received January 2017; accepted March 2017

Abstract. *Traditional heartbeat multi-classification of support vector machine (SVM) has problems of using too many SVMs and occupying large memory space during operating. To solve these problems, this paper proposes a heartbeat multi-classification algorithm of hierarchical SVM which adopts separability assessment among several heartbeats categories based on computing kernel space distance. Then we select two classes of heartbeats which are the hardest to separate as training set. The two classes are combined as a new class to involve next class separability comparison. We repeat the above process until all heartbeats are trained. In this paper, experiments are performed by using the most common six categories of heartbeats from MIT-BIH database. The results show that the average classification accuracy of hierarchical SVM is 95.4% and the number of SVMs is five. This method can reduce the number of samples in training set and memory space occupied by storing and computing under high classification accuracy.*
Keywords: Hierarchical support vector machine, Class separability, Model training, Heartbeat classification

1. Introduction.

As the rapid growth of ageing populations, the incidence of cardiovascular disease is increasing quite significantly. So the diagnosis of cardiovascular disease is such an important way to protect health for the aged. Electrocardiograph (ECG) is one of the most common tools in diagnosing cardiovascular disease in modern times [1]. The automatic recognition of heartbeat is a significant technical method for auxiliary diagnosis. It is also a research hotspot in the field of pattern recognition, medical signal processing and many others [2].

Heartbeat automatic recognition includes two basic steps generally. Firstly, it mines useful information from ECG beats or fragments, and then a heartbeat classifier is established based on the features extracted from the useful information. Eventually, the beat category can be labeled and identification is completed [3]. SVM is a multi-classification extension technique based on binary classifier [4,5]. Due to the fact that beat classification is a complex multi-classification problem, the technique determines calculation and classification accuracy directly [6,7]. At present, there are mainly two methods to construct multi-classification SVM: "one-against-one algorithm" and "one-against-all algorithm" [8].

One-against-all (OAA) algorithm is used for SVM multi-classification. The common process is to construct a series of binary classification SVMs. One class and all the rest classes are selected to classify by each SVM. Then the final judgment is based on the results of all binary classification SVMs [9]. OAA method needs to build k classifiers for k classification problems. When constructing the ith classifier, training samples are labeled as positive that belongs to the ith class and the rest of samples are labeled as negative. In testing process, the corresponding category of maximum decision function value of each binary classifier is chosen as test data category. Another common multi-classification method of SVM is one-against-one (OAO) algorithm. Randomly choose two different categories from k classification problems as a sub-classifier until every category is involved in building SVM sub-classifiers. Thus, there will be $k(k-1)/2$ sub-classifiers. When training the SVM sub-classifier of class M and class N, all samples of class M and class N are respectively marked as positive and negative. In testing process, test data are input into all SVM sub-classifiers and classified into one class, and then the score of this class is plus one. Finally, the highest score class is selected as test data class.

Hierarchical SVM hierarchically rebuilds SVM multi-classification model in the form of binary tree. It retains the advantage of SVM on handling binary classification problems and reduces the total number of SVMs. In view of this, we put forward the heartbeat multi-classification algorithm of hierarchical SVM that each binary tree node is determined based on calculating the kernel spaces distance of different beats categories. The system can reduce error accumulation to ensure heartbeat classification accuracy.

2. **Hierarchical Support Vector Machine.**

The method of hierarchical SVM multi-classification uses binary tree structure to build SVM multi-classifier. The method divides all samples into two classes and then divides each class into two sub-classes, so it goes on until getting a separate category [10]. When classifying, the hierarchical SVM structure does not need to traverse all classifiers and does not show "reject identification area". There are mainly two kinds of binary tree topological structures [11]: complete binary tree structure and partial binary tree structure. They are shown as Figure 1.

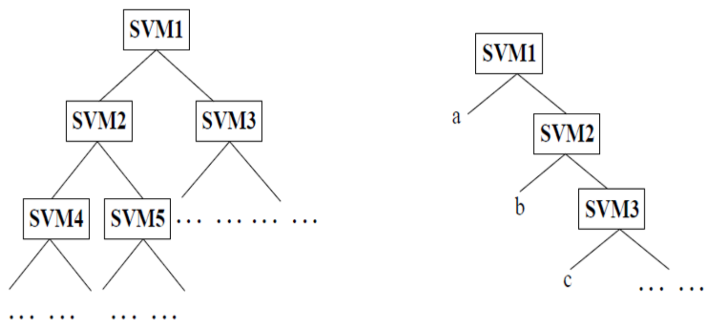


FIGURE 1. Complete and partial binary tree structures

The left is complete binary tree that divides the classes contained at each decision node into two sub-classes with the same quantity. The right is partial binary tree that separates one class from the rest of classes at each decision node. We can see that different building orders for binary tree and the corresponding segmentation of each class are also different. So the generalization performance of classification model is greatly influenced by the upper node classification performance. There are two common ways to generate binary tree: (1) separating the easiest class out, and making the upper SVM sub-classifier have higher generalization performance and the system have lower misclassification rate; (2) separating the most widely distributed class out, and making it have larger segmentation in classification model. The category division order is determined in turn with fully considering distribution characteristics of heartbeat samples and measurement of class separability in feature space.

3. Heartbeat Multi-Classification Algorithm Based on Hierarchical SVM.

3.1. Training model process based on kernel space distance metric. In design of hierarchical SVM structure, class separability is the basis for dividing sub-classifiers. The SVM algorithm design is only for solving binary classification problem; however, heartbeat classification involves a variety of beats categories in practice. As features of ECG signals are linear inseparable data set and high dimension, the distance of kernel spaces is measured as class separability in the paper. So the distance metric is also regarded as the measurements of difference, separating degree and identification capability to reflect class separability fully.

Firstly, let n types of heartbeat matrices $A_1, A_2, \ldots, A_n$ be the input, and heartbeat vectors of each matrix represent the same heartbeat category. In turn, select two heartbeat matrices A_s and A_t with different categories, and the matrix forms are as $\{x_1, x_2, \ldots, x_{n_s}\}$, $\left\{x_1^*, x_2^*, \ldots, x_{n_t}^*\right\}$. The centers of two heartbeat matrices will change after kernel mapping and new centers are m_Φ^s and m_Φ^t. Distance measurement $d^H\left(m_\Phi^s, m_\Phi^t\right)$ between them can be defined as follows:

$$d^H\left(m_\Phi^s, m_\Phi^t\right) = \sqrt{\frac{1}{n_s^2}\sum_{i=1}^{n_s}\sum_{j=1}^{n_s} K(x_i, x_j) - \frac{1}{n_s n_t}\sum_{i=1}^{n_s}\sum_{j=1}^{n_t} K\left(x_i, x_j^*\right) + \frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t} K\left(x_i^*, x_j^*\right)} \tag{1}$$

where n_s, n_t are the number of beats in two heartbeat matrices respectively. $K(\bullet, \bullet)$ is kernel function selected Gauss radial basis kernel function in this paper.

Distance measurement $d^H(x, m_\Phi^i)$ between one heartbeat vector x of matrix and its class center m_Φ^i can be computed as follows:

$$d^H\left(x, m_\Phi^i\right) = \sqrt{K(x, x) - \frac{2}{n}\sum_{i=1}^{n} K(x, x_i) + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} K(x_i, x_j)} \tag{2}$$

Then the standard deviation σ^H of this class of heartbeat is:

$$\sigma^H = \frac{1}{n-1}\sum_{i=1}^{n} d^H\left(x_i, m_\Phi^i\right) \tag{3}$$

At last, the class separability rsm_{st} between two categories of heartbeats can be expressed by distance measurement $d^H(x, m_\Phi^i)$ for different classes after mapping in new feature space from Equation (2) and class standard deviation σ^H from Equation (3). It is defined as:

$$rsm_{st} = \frac{d^H\left(m_\Phi^s, m_\Phi^t\right)}{\left(\sigma_s^H + \sigma_t^H\right)} \tag{4}$$

Figure 2 presents the training process of hierarchical SVM model. According to class separability rsm_{st}, the paper utilizes the proposed algorithm to train the model.

Selecting heartbeat training set firstly, in which K classes of samples are initialized to K leaf nodes. Judging the number of clusters, if $K > 2$, class separability rsm_{st} will be calculated for K classes of heartbeat matrices to evaluate separability for any two classes. Comparing and ranking the distance metric values rsm_{st} based on kernel space, two corresponding clusters M and N with minimum values are selected. Then set threshold value of misclassification cost to improve classification accuracy. Class M will be labeled as positive and class N will be as negative when training single SVM binary classifier. At the same time, the number of hierarchical SVM layers is one. Classes M and N will be combined as a new class for subsequent measurement comparison. So the number of clusters K reduces one. Repeat the above process until all heartbeats are involved in training SVM. When $K = 2$, the SVM sub-classifier in the top layer is trained to output two categories and the training process finishes.

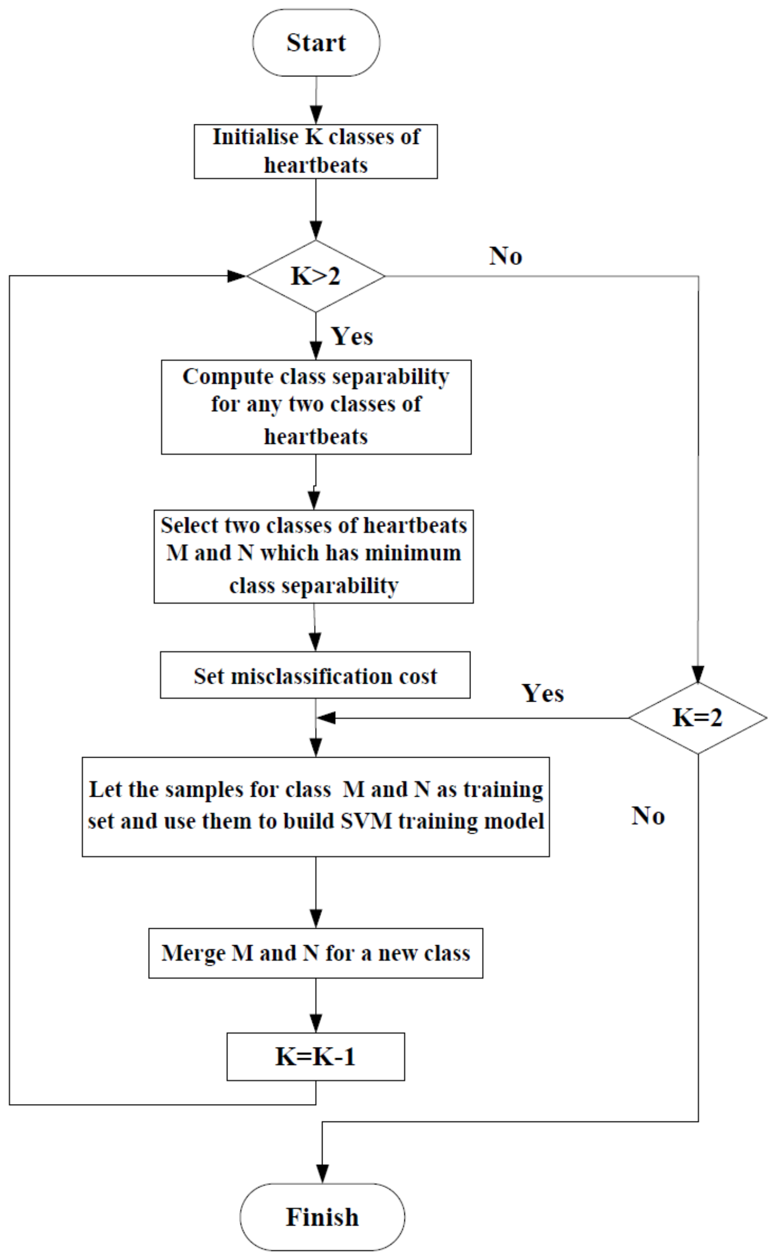


FIGURE 2. Hierarchical SVM model training flowchart

3.2. **Classification process of the proposed heartbeat multi-classification algorithm.** The paper selects six classes of common arrhythmia heartbeats as input for experiments. These are normal beat (N), left bundle branch block beat (L), right bundle branch block beat (R), paced beat (/), atrial premature beat (A) and premature ventricular contraction (V). The training set selection adopts random selection method from pure heartbeats after preprocessing. So the training set not only contains all categories of heartbeats, but also the proportion of the number of each categories in training set is the same as the proportion in original sample set. Table 1 presents the class separability values for any two classes of heartbeats among six classes.

As shown in Table 1, the value of class separability between L cluster and V cluster is minimum, so the two classes have the worst separability. According to hierarchical SVM theory, place the two classes on the bottom of hierarchical SVM model which can build SVM1. Then we merge the two beats clusters as a new beats cluster, and compute class

Table 1. Class separability for any two classes of heartbeats among six classes

	L	**R**	**N**	**A**	**V**	**/**
L	–	0.6623	0.6447	0.7793	***0.4882***	0.6857
R	0.6623	–	0.6807	0.8432	0.5571	0.7536
N	0.6447	0.6807	–	0.8333	0.5476	0.7511
A	0.7793	0.8432	0.8333	–	0.6365	0.8657
V	***0.4882***	0.5571	0.5476	0.6365	–	0.5544
/	0.6857	0.7536	0.7511	0.8657	0.5544	–

Table 2. Class separability among five classes

	L+V	**R**	**N**	**A**	**/**
L+V	–	0.5867	***0.5722***	0.6805	0.5967
R	0.5867	–	0.6807	0.8432	0.7536
N	***0.5722***	0.6807	–	0.8333	0.7511
A	0.6805	0.8432	0.8333	–	0.8657
/	0.5967	0.7536	0.7511	0.8657	–

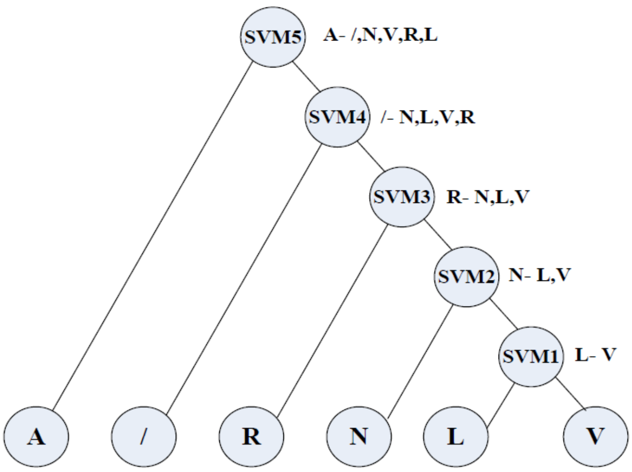


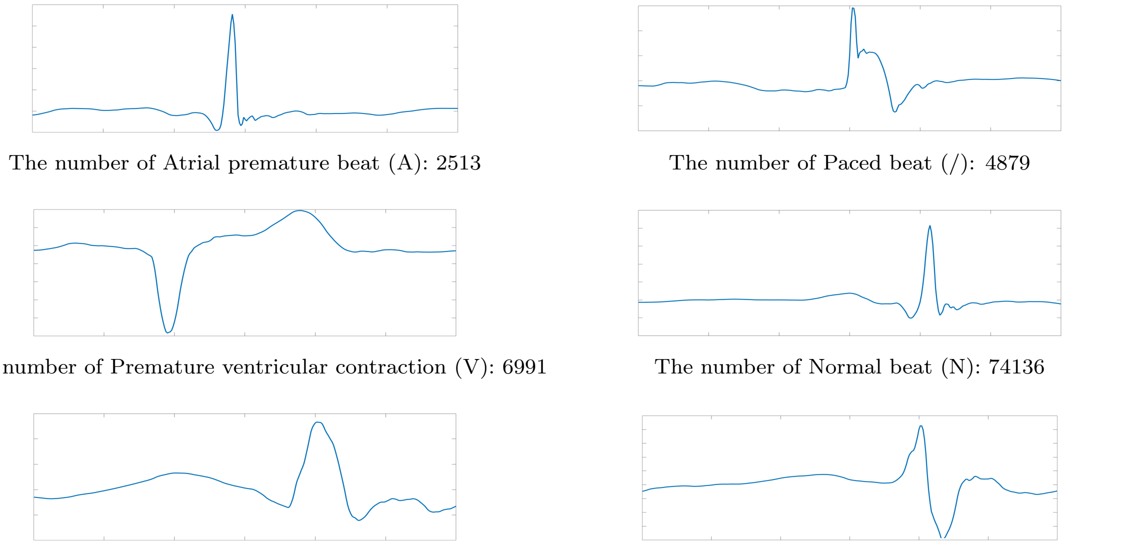
Figure 3. Hierarchical SVM model of six classes' heartbeats

separability between the new cluster and the rest clusters. The results are shown in Table 2.

From Table 2, L+V expresses the new beats cluster after merging L cluster and V cluster. According to the results in Table 2, value of class separability between new beats cluster L+V and N cluster is minimum, so the separability is poor. Let the two classes be the second layer from bottom in which hierarchical SVM model can build SVM2. Then we merge the two clusters as a new cluster and compute in the same way as above. The final multi-classification model based on hierarchical SVM is shown in Figure 3.

4. **Experimental Results and Analysis.** We evaluate our approach on MIT-BIH ECG database [12], and the samples used include normal beats and five categories of arrhythmia beats which are relatively common categories and have a certain scale of quantity. According to [13], these kinds of heartbeats in MIT-BIH Arrhythmia Database are typical that could predict occurrence of some heart diseases. The selected number of heartbeats is shown in Figure 4.

In experiments of verifying validity of classifier, choose time-domain feature of heartbeats as SVM input. The method of heartbeat segmentation is according to [14], getting 300 dimensional time-domain feature after resampling. The number of samples is selected



The number of Atrial premature beat (A): 2513

The number of Paced beat (/): 4879

The number of Premature ventricular contraction (V): 6991

The number of Normal beat (N): 74136

The number of Right bundle branch block beat (R): 7172

The number of Left bundle branch block beat (L): 7971

FIGURE 4. Six classes of heartbeats and quantity

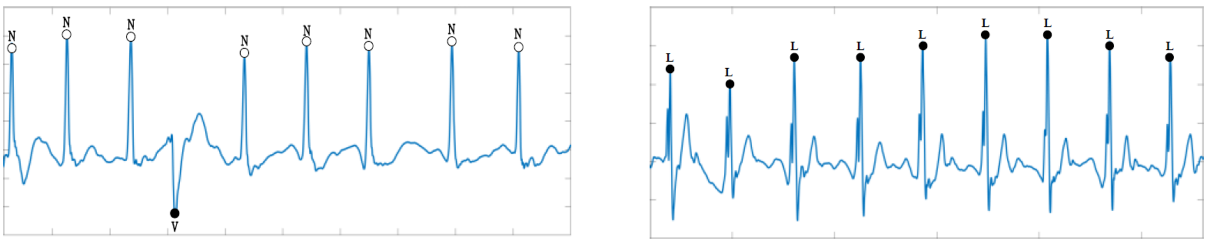


FIGURE 5. No.210 and No.111 ECG signal segments classification results

TABLE 3. Classification accuracy for six classes based on hierarchical SVM

Class	**N**	**A**	**V**	**/**	**L**	**R**
Total number	74136	2513	6991	4879	7971	7172
The number of correct classification	73433	1327	5339	4767	7283	6741
Classification accuracy	99.05%	52.80%	76.36%	97.70%	91.36%	93.99%

of 1200 as training set and the rest of heartbeats are used as testing set to verify algorithm availability. Taking segment of No.210 and No.111 signals as example, classification results are shown in Figure 5.

Just as shown in Figure 5, detections of abnormal heartbeats are marked as solid points and N beats are as hollow points. The abnormal heartbeats are L beats and V beats. The final classification results for six categories of heartbeats are shown as Table 3.

As shown in Table 3, our method can effectively identify N beats, / beats, L beats and R beats. As waveforms of A beats and V beats are complex and changeable, simple time-domain features only adopted in this paper are difficult to reflect signal difference. So the recognition accuracies of those are relatively low. The average classification accuracy is 95.4% which is the ratio of the number of correct classification to the total number of heartbeats for six classes. [8] respectively utilizes OAO and OAA algorithms to identify heartbeat categories. The comparison results of experiments are shown in Table 4.

From Table 4 we know that the classification accuracy of our approach is not inferior to traditional algorithm. All experiments could almost identify all the N beats. It also has high recognition accuracy which is more than 90% for L beats, R beats and / beats. However, to V beats, recognition accuracy of our model is higher than traditional model

Table 4. Comparison for classification accuracy based on our method and traditional algorithms

	N	L	V
OAO algorithm [8]	100%	85.71%	70%
OAA algorithm [8]	100%	92.85%	70%
The proposed algorithm	99.05%	91.36%	76.36%

Table 5. The number of SVMs needed by all kinds of multi-classification algorithms

	The proposed model	OAO model	OAA model
The number of needed SVMs	5	15	6

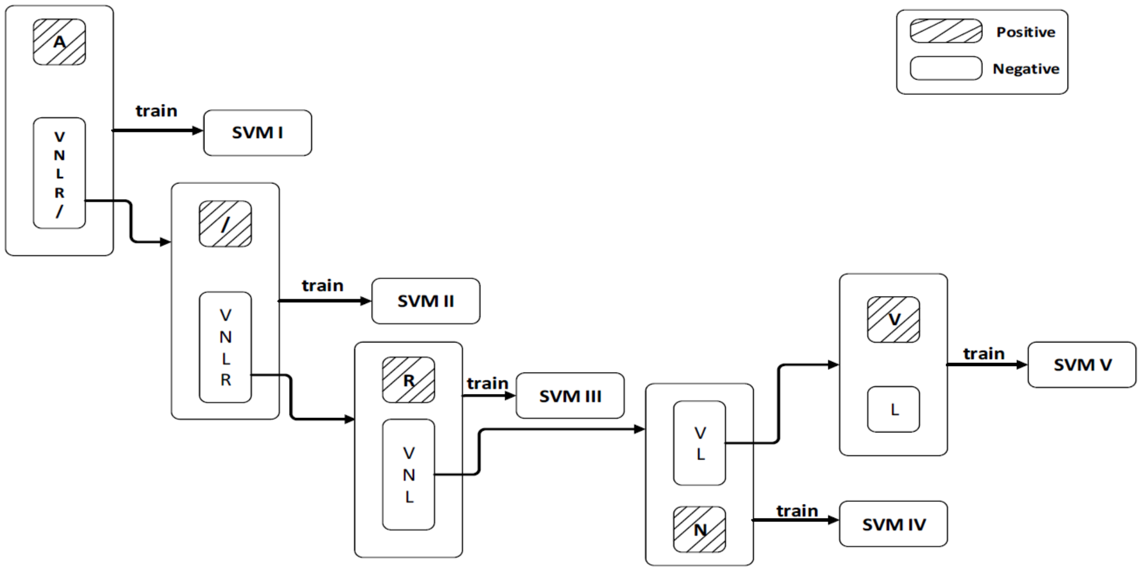


Figure 6. Variations chart for training set in the proposed model training process

that only uses time-domain features. The number of SVMs needed for classifying six classes of heartbeats is compared in Table 5.

As shown in Table 5, the number of SVMs for the proposed model is much less than OAO algorithm model. Though it has only one less number of SVM compared with OAA model, all samples are commonly adopted by OAA algorithm when training each SVM. Thus, a lot of memory space and time in storage and calculation process could be occupied. However, training sample quantity of each layer for the proposed model presents a trend of decreasing layer-by-layer. The samples change of hierarchical SVM in training process is shown in Figure 6.

As shown in Figure 6, when training each binary classification of SVM, positive samples used in upper layer are abandoned. So training samples and time will decrease layer-by-layer, and the memory in operation will be effectively reduced.

5. **Conclusion.** In order to solve problems of traditional heartbeat classification algorithms, this paper presents a heartbeat multi-classification algorithm of hierarchical SVM which builds a bottom-up model based on heartbeat distance measurement values of kernel space. The classification process is according to the order from top to bottom. Experimental results show that the number of SVMs needed is five when classifying six categories of heartbeats. And samples of training set reduce from top to bottom layer-by-layer. Compared with traditional SVM classification model, this model not only ensures high accuracy rate, but also greatly reduces the number of SVMs and shortens operation

time. So the method is so helpful for ECG automatic recognition that can effectively help doctors diagnose and treat various heart diseases for ageing people. However, this paper only adopts heartbeat time-domain characteristics for experiments, and the recognition accuracy of two kinds of complex and changeable waveforms is not very high. In future research, we will change input characteristics types of heartbeats to make classification results perfect.

Acknowledgment. We thank all the volunteers and colleagues provided helpful comments on previous versions of the manuscript. This work was supported by Zhuhai Advantage Discipline Construction Project and International Joint Center for Welfare of Zhuhai College of Jilin University. This work was also supported by the Science and Technology Development projects funded by Jilin Government (20150204039GX, 20170414017GH), Science and Technology Development Special Funded by Guangdong Government (2016A030313658), the Key Scientific and Technological Special Fund Project supported by Changchun Government under Grant No. 14KG064, and the National Natural Science Foundation of China under Grant No. 61572228.

REFERENCES

[1] Z. Zidelmal et al., ECG beat classification using a cost sensitive classifier, *Computer Methods & Programs in Biomedicine*, vol.111, no.3, pp.570-577, 2013.

[2] D. Saibal, A. Chatterjee and S. Munshi, Correlation technique and least square support vector machine combine for frequency domain based ECG beat classification, *Medical Engineering & Physics*, vol.32, no.10, pp.1161-1169, 2010.

[3] M. A. Haque Monil et al., Bangla text processing and recognition based on fuzzy unsupervised feature extraction and SVM, *IEEE International Conference on Machine Learning and Cybernetics*, vol.3, pp.1272-1278, 2013.

[4] T. Liu et al., Electrocardiogram beat classification based on kernel principal component analysis and support vector machine, *Optics and Precision Engineering*, vol.23, no.10z, pp.744-751, 2015.

[5] D. Wang et al., Heart beat classification based on feature fusion by principle component analysis, *Optics and Precision Engineering*, vol.23, no.10z, pp.453-458, 2015.

[6] T. Liu et al., Vector quantization for ECG beats classification, *IEEE International Conference on Computational Science and Engineering*, pp.13-20, 2014.

[7] C.-P. Shen et al. Detection of cardiac arrhythmia in electrocardiograms using adaptive feature extraction and modified support vector machines, *Expert Systems with Applications: An International Journal*, vol.39, no.9, pp.7845-7852, 2012.

[8] M. Kallas et al., Multi-class SVM classification combined with kernel PCA feature extraction of ECG signals, *IEEE International Conference on Telecommunications*, pp.1-5, 2012.

[9] M. Hejazi et al., Multiclass support vector machines for classification of ECG data with missing values, *Applied Artificial Intelligence*, vol.29, no.7, pp.660-674, 2015.

[10] B. Wei and J. T. Kwok, Multi-label classification on tree- and DAG-structured hierarchies, *International Conference on Machine Learning*, pp.17-24, 2011.

[11] C. Bernhard-Oettel, M. Sverke and H. D. Witte, Multi-label classification by mining label and instance correlations from heterogeneous information networks, *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol.1703, no.2, pp.614-622, 2013.

[12] R. Mark and G. Mody, *MIT-BIH Arrhythmia Database*, http//:www.physionet.org/, 1997.

[13] S. N. Yu and Y. H. Chen, Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network, *Pattern Recognition Letters*, vol.28, no.10, pp.1142-1150, 2007.

[14] E. M. Imah, W. Jatmiko and T. Basaruddin, Adaptive multilayer generalized learning vector quantization (AMGLVQ) as new algorithm with integrating feature extraction and classification for arrhythmia heartbeats classification, *IEEE International Conference on Systems, Man, and Cybernetics*, vol.2, pp.150-155, 2012.